



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/808,073	03/24/2004	Xiaoyun Zhu	200313904-1	5369

22879 7590 11/13/2009
HEWLETT-PACKARD COMPANY
Intellectual Property Administration
3404 E. Harmony Road
Mail Stop 35
FORT COLLINS, CO 80528

EXAMINER

TANG, KENNETH

ART UNIT	PAPER NUMBER
----------	--------------

2195

NOTIFICATION DATE	DELIVERY MODE
-------------------	---------------

11/13/2009

ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

JERRY.SHORMA@HP.COM
ipa.mail@hp.com
laura.m.clark@hp.com

Office Action Summary	Application No. 10/808,073	Applicant(s) ZHU ET AL.	
	Examiner KENNETH TANG	Art Unit 2195	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 15 October 2009 and 06 July 2009.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1 and 4-27 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1 and 4-27 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☒ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____ |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date <u>10/15/09</u> . | 6) <input type="checkbox"/> Other: _____ |

Art Unit: 2195

DETAILED ACTION

1. Claims 1 and 4-27 are presented for examination.
2. This action is in response to the Amendment/Response on 7/6/09 and the IDS submission on 10/15/09. Applicant's arguments have been fully considered but are moot in view of the new grounds of rejections.

Specification

3. The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed.

Information Disclosure Statement

4. The IDS on 10/15/09 has been considered and initialed by the Examiner.

Double Patenting

The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent and to prevent possible harassment by multiple assignees. A nonstatutory obviousness-type double patenting rejection is appropriate where the conflicting claims are not identical, but at least one examined application claim is not patentably distinct from the reference claim(s) because the examined application claim is either anticipated by, or would have been obvious over, the reference claim(s). See, e.g., *In re Berg*, 140 F.3d 1428, 46 USPQ2d 1226 (Fed. Cir. 1998); *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); *In re Van Ornum*, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); *In re Vogel*, 422 F.2d 438, 164 USPQ 619 (CCPA 1970); and *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) or 1.321(d) may be used to overcome an actual or provisional rejection based on a nonstatutory double patenting ground provided the conflicting application or patent either is shown to be commonly owned

Art Unit: 2195

with this application, or claims an invention made as a result of activities undertaken within the scope of a joint research agreement.

Effective January 1, 1994, a registered attorney or agent of record may sign a terminal disclaimer. A terminal disclaimer signed by the assignee must fully comply with 37 CFR 3.73(b).

5. Claims 1 and 7 are rejected on the ground of nonstatutory obviousness-type double patenting as being unpatentable over claims 1, 3, and 4 of U.S. Patent No. 7,146,353 B2 in view of Hill et al. (US 2004/0267897 A1).

6. The following table mapping illustrates the limitations of the Instant Application that are disclosed in the claims of US 7,146,353 B2 (bolded emphasis by Examiner):

Instant Application	US 7,146,353 B2
<p>1. A processor-based method for allocating resources to a plurality of applications, comprising:</p> <p>determining available resources of a networked computing system;</p>	<p>1. A processor-implemented method for allocating resources to a plurality of applications, wherein the resources include a plurality of servers and at least one of the applications uses a tiered arrangement of servers, comprising:</p> <p>gathering instrumentation data for work requests processed by the applications;</p> <p><i>(disclosed in the reference of Hill)</i></p>

determining, for each application, required resources of the application;	<p>determining an associated workload level for work requests processed by the applications;</p> <p>determining for each application a first application resource requirement as a function of the workload levels and a service level metric associated with the application; representing each server as a processor-sharing queue having at least one critical resource;</p> <p>determining respective average response times of each of the tiers, each respective average response time being a function of a number servers in the tier, an arrival rate of work requests, and an average utilization rate of the critical resource;</p> <p>determining a total average response time as a sum of the respective average response times of each of the tiers;</p> <p>determining a minimum total number of servers required in each tier for the total</p>
--	--

<p>determining an assigned subset of the available resources for each application as a function of the required resources of the application and the available resources, wherein the function reduces communication delays between resources of the subset of the available resources in conformance with bandwidth capacity requirements of the application and in conformance with network bandwidth limitations; and</p> <p>associating the applications with the assigned subsets of resources.</p>	<p>average response time of the application to satisfy the service level metric;</p> <p>determining for each application an assigned subset of resources as a function of the first application resource requirement, wherein the function minimizes communication delays between resources, and satisfies a bandwidth capacity requirement of the application; and</p> <p><i>(reference of Hill teaches the missing limitations)</i></p> <p>automatically reconfiguring the resources consistent with the assigned subset of resources for each application.</p> <p>3. The method of claim 1, wherein the step</p>
---	---

<p>7. The method of claim 1, wherein reducing the communications delay between resources comprises solving a mixed-integer programming problem.</p>	<p>of determining an assigned subset of resources comprises assigning resources to tiers by a function that satisfies the resource requirements associated with each tier and minimizes communication delay between servers.</p> <p>4. The method of claim 3, wherein the function is a mixed-integer programming function.</p>
--	--

7. As illustrated above, the bolded portions of the claimed Instant Application are disclosed in the corresponding bolded portions of US 7,146,353 B2. US 7,146,353 B2 does not disclose the claimed limitations of and “determining available resources of a networked computing system;” and using said available resources as part of the function that also is “in conformance with network bandwidth limitations”, as illustrated in the unbolded portion under the column representing the Instant Application. However, Hill discloses a processor-based method for allocating resources to a plurality of applications (see Abstract), comprises determining available resources of a networked computing system (lines 7-8 of [0023]); determining, for each application, required resources of the application (lines 9-17 of [0023]); determining an assigned

Art Unit: 2195

subset of the available resources for each application as a function of the required resources of the application and the available resources, wherein the function reduces communication delays between resources of the subset of the available resources in conformance with bandwidth capacity requirements of the application and in conformance with network bandwidth limitations ([0023], [0207], [0208], [0210], [0067]); and associating the applications with the assigned subsets of resources ([0207], [0208], [0177]). One of ordinary skill in the art would have known to modify the claimed invention of US 7,146,353 B2 such that it would determine available resources of a networked computing system and use said available resources as part of the function that also is in conformance with network bandwidth limitations. The suggestion/motivation for doing so would have been to better manage the resource requirements of application running across the plurality of resources, to balancing the load between them, to more efficiently conserve bandwidth, thus, improving the overall system ([0017], [0020]).

8. US 7,146,353 B2 and Hill do not expressly disclose that its application resource optimization ([0093]) is based on an objective function to reduce a number of network hops between processing resources in the assigned subset.

9. However, Lee teaches allocating network resources along an optimal path, wherein the path selection within routing is typically formulated as a shortest path optimization problem, i.e., by determining a series of network links connecting the source and destination such that a particular objective function is minimized. The objective function may be the number of hops, cost, delay, or some other metric (see Abstract; page 46, 1st two paragraphs). Hill and Lee are analogous art because they are both in the same field of endeavor of network resource allocation and both are attempting to optimize its resource allocation such that delay is reduced. Thus, one

of ordinary skill in the art would have known to modify Hill's network resource allocation system such that it would utilize an objective function to reduce a number of network hops between processing resources in the assigned subset, as taught in the reference of Lee. The suggestion/motivation for doing so would have been to provide an effective scheme to specifically calculate and determine the optimization of a path involving the minimal number of hops, delays, costs, etc. (see Abstract; page 46, 1st two paragraphs).

10. Claim 9 is rejected on the ground of nonstatutory obviousness-type double patenting as being unpatentable over claim 10 of U.S. Patent No. 7,146,353 B2 in view of Hill et al. (US 2004/0267897 A1).

11. The following table mapping illustrates the limitations of the Instant Application that are disclosed in the claims of US 7,146,353 B2 (bolded emphasis by Examiner):

Instant Application	US 7,146,353 B2
<p>9. A system comprising:</p> <p>means for determining available resources</p>	<p>10. An apparatus for allocating resources to a plurality of applications, wherein the resources include a plurality of servers and at least one of the applications uses a tiered arrangement of servers, comprising:</p> <p><i>(disclosed in the reference of Hill)</i></p>

Art Unit: 2195

<p>of a networked computing system;</p> <p>means for determining required resources for each application of a plurality of applications;</p>	<p>means for gathering instrumentation data for work requests processed by the applications;</p> <p>means for determining an associated workload level for work requests processed by the applications;</p> <p>means for generating for each application a first application resource requirement as a function of the workload levels and a service level metric associated with the application;</p> <p>means for representing each server as a processor-sharing queue having at least one critical resource; means for determining respective average response times of each of the tiers, each respective average response time being a function of a number servers in the tier, an arrival rate of work requests, and an average utilization rate of the critical resource;</p>
---	---

<p>means for determining an assigned subset of the available resources for each application as a function of the required resources of the application and the available resources, wherein the function reduces communication delays between resources of the subset of the available resources in conformance with bandwidth capacity requirements of the application and in conformance with network bandwidth limitations; and</p> <p>means for associating the applications with the assigned subsets of resources.</p>	<p>means for determining a total average response time as a sum of the respective average response times of each of the tiers;</p> <p>means for determining a minimum total number of servers required in each tier for the total average response time of the application to satisfy the service level metric;</p> <p>means for determining for each application an assigned subset of resources as a function of the first application resource requirement, wherein the function minimizes communication delays between resources, and satisfies a bandwidth capacity requirement of the application; and</p> <p><i>(reference of Hill teaches the missing limitations)</i></p> <p>means for automatically reconfiguring the resources consistent with the assigned subset of resources for each application.</p>
---	---

Art Unit: 2195

12. As illustrated above, the bolded portions of the claimed Instant Application are disclosed in the corresponding bolded portions of US 7,146,353 B2. US 7,146,353 B2 does not disclose the claimed limitations of and “determining available resources of a networked computing system;”, “determining” the required resources, and using said available resources as part of the function that also is “in conformance with network bandwidth limitations”, as illustrated in the unbolded portion under the column representing the Instant Application. However, Hill discloses a processor-based method for allocating resources to a plurality of applications (see Abstract), comprises determining available resources of a networked computing system (lines 7-8 of [0023]); determining, for each application, required resources of the application (lines 9-17 of [0023]); determining an assigned subset of the available resources for each application as a function of the required resources of the application and the available resources, wherein the function reduces communication delays between resources of the subset of the available resources in conformance with bandwidth capacity requirements of the application and in conformance with network bandwidth limitations ([0023], [0207], [0208], [0210], [0067]); and associating the applications with the assigned subsets of resources ([0207], [0208], [0177]). One of ordinary skill in the art would have known to modify the claimed invention of US 7,146,353 B2 such that it would determine and take into consideration the available resources of the networked computer system, “determining” the required resources, and being in conformance with network bandwidth limitations. The suggestion/motivation for doing so would have been to better manage the resource requirements of application running across the plurality of resources, to balancing the load between them, to more efficiently conserve bandwidth, thus, improving the overall system ([0017], [0020]).

Art Unit: 2195

13. US 7,146,353 B2 and Hill do not expressly disclose that its application resource optimization ([0093]) is based on an objective function to reduce a number of network hops between processing resources in the assigned subset.

14. However, Lee teaches allocating network resources along an optimal path, wherein the path selection within routing is typically formulated as a shortest path optimization problem, i.e., by determining a series of network links connecting the source and destination such that a particular objective function is minimized. The objective function may be the number of hops, cost, delay, or some other metric (see Abstract; page 46, 1st two paragraphs). Hill and Lee are analogous art because they are both in the same field of endeavor of network resource allocation and both are attempting to optimize its resource allocation such that delay is reduced. Thus, one of ordinary skill in the art would have known to modify Hill's network resource allocation system such that it would utilize an objective function to reduce a number of network hops between processing resources in the assigned subset, as taught in the reference of Lee. The suggestion/motivation for doing so would have been to provide an effective scheme to specifically calculate and determine the optimization of a path involving the minimal number of hops, delays, costs, etc. (see Abstract; page 46, 1st two paragraphs).

15. Claims 10 and 16 are rejected on the ground of nonstatutory obviousness-type double patenting as being unpatentable over claims 13 and 16 of U.S. Patent No. 7,146,353 B2 in view of Hill et al. (US 2004/0267897 A1).

16. The following table mapping illustrates the limitations of the Instant Application that are disclosed in the claims of US 7,146,353 B2 (bolded emphasis by Examiner):

Instant Application	US 7,146,353 B2
<p>10. A computer-readable medium configured with instructions for causing a processor of a data processing arrangement to allocate resources to a plurality of applications, comprising:</p> <p>determining available resources of a networked computing system;</p>	<p>13. An article of manufacture for allocating resources to a plurality of applications, wherein the resources include a plurality of servers and at least one of the applications uses a tiered arrangement of servers, comprising:</p> <p>a computer-readable medium configured with instructions for causing a processor-based system to perform the steps of, gathering instrumentation data for work requests processed by the applications;</p> <p><i>(disclosed in the reference of Hill)</i></p> <p>determining an associated workload level for work requests processed by the applications;</p>

<p>determining, for each application, required resources of the application;</p>	<p>generating for each application a first application resource requirement as a function of the workload levels and a service level metric associated with the application;</p> <p>representing each server as a processor-sharing queue having at least one critical resource;</p> <p>determining respective average response times of each of the tiers, each respective average response time being a function of a number servers in the tier, an arrival rate of work requests, and an average utilization rate of the critical resource;</p> <p>determining a total average response time as a sum of the respective average response times of each of the tiers;</p> <p>determining a minimum total number of servers required in each tier for the total average response time of the application to satisfy the service level metric;</p>
---	---

<p>determining an assigned subset of the available resources for each application as a function of the required resources of the application and the available resources, wherein the function reduces communication delays between resources of the subset of the available resources in conformance with bandwidth capacity requirements of the application and in conformance with network bandwidth limitations; and</p> <p>associating the applications with the assigned subsets of resources.</p>	<p>determining for each application an assigned subset of resources as a function of the first application resource requirement, wherein the function minimizes communication delays between resources, and satisfies a bandwidth capacity requirement of the application; and</p> <p><i>(the reference of Hill teaches the missing limitations)</i></p> <p>automatically reconfiguring the resources consistent with the assigned subset of resources for each application.</p> <p>15. The article of manufacture of claim 13, wherein the computer-readable medium is further configured with instructions for causing a processor-based system to, in determining an assigned subset of resources, perform the step</p>
--	--

16. The computer-readable medium of claim 10, wherein reducing the communications delay between resources comprises solving a mixed-integer programming problem.	<p>of assigning resources to tiers by a function that satisfies the resource requirements associated with each tier and minimizes communication delay between servers.</p> <p>16. The article of manufacture of claim 15, wherein the function is a mixed-integer programming function.</p>
---	--

17. As illustrated above, the bolded portions of the claimed Instant Application are disclosed in the corresponding bolded portions of US 7,146,353 B2. US 7,146,353 B2 does not disclose the claimed limitations of and “determining available resources of a networked computing system;”, “determining” the required resources, and using said available resources as part of the function that also is “in conformance with network bandwidth limitations”, as illustrated in the unbolded portion under the column representing the Instant Application. However, Hill discloses a processor-based method for allocating resources to a plurality of applications (see Abstract), comprises determining available resources of a networked computing system (lines 7-8 of [0023]); determining, for each application, required resources of the application (lines 9-17 of [0023]); determining an assigned subset of the available resources for each application as a

Art Unit: 2195

function of the required resources of the application and the available resources, wherein the function reduces communication delays between resources of the subset of the available resources in conformance with bandwidth capacity requirements of the application and in conformance with network bandwidth limitations ([0023], [0207], [0208], [0210], [0067]); and associating the applications with the assigned subsets of resources ([0207], [0208], [0177]). One of ordinary skill in the art would have known to modify the claimed invention of US 7,146,353 B2 such that it would determine and take into consideration the available resources of the networked computer system, “determining” the required resources, and being in conformance with network bandwidth limitations. The suggestion/motivation for doing so would have been to better manage the resource requirements of application running across the plurality of resources, to balancing the load between them, to more efficiently conserve bandwidth, thus, improving the overall system ([0017], [0020]).

18. US 7,146,353 B2 and Hill do not expressly disclose that its application resource optimization ([0093]) is based on an objective function to reduce a number of network hops between processing resources in the assigned subset.

19. However, Lee teaches allocating network resources along an optimal path, wherein the path selection within routing is typically formulated as a shortest path optimization problem, i.e., by determining a series of network links connecting the source and destination such that a particular objective function is minimized. The objective function may be the number of hops, cost, delay, or some other metric (see Abstract; page 46, 1st two paragraphs). Hill and Lee are analogous art because they are both in the same field of endeavor of network resource allocation and both are attempting to optimize its resource allocation such that delay is reduced. Thus, one

Art Unit: 2195

of ordinary skill in the art would have known to modify Hill's network resource allocation system such that it would utilize an objective function to reduce a number of network hops between processing resources in the assigned subset, as taught in the reference of Lee. The suggestion/motivation for doing so would have been to provide an effective scheme to specifically calculate and determine the optimization of a path involving the minimal number of hops, delays, costs, etc. (see Abstract; page 46, 1st two paragraphs).

20. Claims 18 and 22 are rejected on the ground of nonstatutory obviousness-type double patenting as being unpatentable over claims 1, 3, and 4 of U.S. Patent No. 7,146,353 B2 in view of Hill et al. (US 2004/0267897 A1).

21. The following table mapping illustrates the limitations of the Instant Application that are disclosed in the claims of US 7,146,353 B2 (bolded emphasis by Examiner):

Instant Application	US 7,146,353 B2
18. A system, comprising: a plurality of network-coupled processing resources; a plurality of storage resources network-coupled to the processing resources, wherein the processing and storage resources are allocated to a plurality of applications;	1. A processor-implemented method for allocating resources to a plurality of applications, wherein the resources include a plurality of servers and at least one of the applications uses a tiered arrangement of servers, comprising:

<p>a computing arrangement configured to, determine, for each application of the plurality of applications, required resources of the application;</p>	<p>gathering instrumentation data for work requests processed by the applications;</p> <p>determining an associated workload level for work requests processed by the applications;</p> <p>determining for each application a first application resource requirement as a function of the workload levels and a service level metric associated with the application;</p> <p>representing each server as a processor-sharing queue having at least one critical resource;</p> <p>determining respective average response times of each of the tiers, each respective average response time being a function of a number servers in the tier, an arrival rate of work requests, and an average utilization rate of the critical resource;</p> <p>determining a total average response time as a sum of the respective average response times of each of the tiers;</p>
---	--

<p>determining an assigned subset of the processing and storage resources for each application as a function of the required resources of the application and the processing and storage resources, wherein the function reduces communication delays between resources of the subset of the network and processing resources in conformance with bandwidth capacity requirements of the application and in conformance with network bandwidth limitations;</p> <p>associate the applications with the assigned subsets of processing and storage resources.</p>	<p>determining a minimum total number of servers required in each tier for the total average response time of the application to satisfy the service level metric;</p> <p>determining for each application an assigned subset of resources as a function of the first application resource requirement, wherein the function minimizes communication delays between resources, and satisfies a bandwidth capacity requirement of the application; and</p> <p><i>(reference of Hill teaches the missing limitations)</i></p> <p>automatically reconfiguring the resources consistent with the assigned subset of resources for each application.</p>
--	---

<p>22. The system of claim 18, wherein the computing arrangement is configured to reduce the communications delay between resources by solving a mixed-integer programming problem.</p>	<p>3. The method of claim 1, wherein the step of determining an assigned subset of resources comprises assigning resources to tiers by a function that satisfies the resource requirements associated with each tier and minimizes communication delay between servers.</p> <p>4. The method of claim 3, wherein the function is a mixed-integer programming function.</p>
--	---

22. As illustrated above, the bolded portions of the claimed Instant Application are disclosed in the corresponding bolded portions of US 7,146,353 B2. US 7,146,353 B2 does not disclose the claimed limitations of a system consisting of a plurality of storage resources network-coupled to the processing resources, wherein the storage resources are allocated to a plurality of applications, in addition to determining an assigned subset of the storage resources for each application as a function of the required resources of the application and the storage resources, wherein the function reduces communication delays between resources of the subset of the

Art Unit: 2195

network and processing resources in conformance with network bandwidth limitations.

However, Hill discloses a processor-based system and method for allocating resources to a plurality of applications (see Abstract), comprises determining available resources of a networked computing system (lines 7-8 of [0023]); determining, for each application, required resources of the application (lines 9-17 of [0023]); determining an assigned subset of the available resources for each application as a function of the required resources of the application and the available resources, wherein the function reduces communication delays between resources of the subset of the available resources in conformance with bandwidth capacity requirements of the application and in conformance with network bandwidth limitations ([0023], [0207], [0208], [0210], [0067]); and associating the applications with the assigned subsets of resources ([0207], [0208], [0177]). Hill also teaches wherein the available resources comprise processing resources, networking resources, and storage resources (Fig. 1, items 101, 102, 103, 115, 116, 111). One of ordinary skill in the art would have known to modify the claimed invention of US 7,146,353 B2 such that it would include a system consisting of a plurality of storage resources network-coupled to the processing resources, wherein the storage resources are allocated to a plurality of applications, in addition to determining an assigned subset of the storage resources for each application as a function of the required resources of the application and the storage resources, wherein the function reduces communication delays between resources of the subset of the network and processing resources in conformance with network bandwidth limitations, as taught in Hill. The suggestion/motivation for doing so would have been to provide a system to perform its method, which better manages the resource requirements

Art Unit: 2195

of applications running across the plurality of resources, balances the load between them, more efficiently conserves bandwidth, and thus, improves the overall system ([0017], [0020]).

23. US 7,146,353 B2 and Hill do not expressly disclose that its application resource optimization ([0093]) is based on an objective function to reduce a number of network hops between processing resources in the assigned subset.

24. However, Lee teaches allocating network resources along an optimal path, wherein the path selection within routing is typically formulated as a shortest path optimization problem, i.e., by determining a series of network links connecting the source and destination such that a particular objective function is minimized. The objective function may be the number of hops, cost, delay, or some other metric (see Abstract; page 46, 1st two paragraphs). Hill and Lee are analogous art because they are both in the same field of endeavor of network resource allocation and both are attempting to optimize its resource allocation such that delay is reduced. Thus, one of ordinary skill in the art would have known to modify Hill's network resource allocation system such that it would utilize an objective function to reduce a number of network hops between processing resources in the assigned subset, as taught in the reference of Lee. The suggestion/motivation for doing so would have been to provide an effective scheme to specifically calculate and determine the optimization of a path involving the minimal number of hops, delays, costs, etc. (see Abstract; page 46, 1st two paragraphs).

Art Unit: 2195

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

25. Claims 1, 4, 9-13, and 18-19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hill et al. (hereinafter Hill) (US 2004/0267897 A1) in view of Lee (“Routing Subject to Quality of Service Constraints in Integrated Communication Networks”, July/August 1995).

26. As to claim 1, Hill teaches a method for allocating resources to a plurality of applications (see Abstract), comprising:

determining, by a computer, available resources of a networked computing system, wherein the available resources comprise processing resources, networking resources, and storage resources (lines 7-8 of [0023]; Fig. 1, items 101, 102, 103, 115, 116, 111; lines 3-6 of paragraph [0097]);

determining, by the computer, for each application, required resources of the application (lines 9-17 of [0023]);

determining, by the computer, an assigned subset of the available resources for each application as a function of the required resources of the application and the available resources, wherein the function reduces communication delays between resources of the subset of the

Art Unit: 2195

available resources in conformance with bandwidth capacity requirements of the application and in conformance with network bandwidth limitations ([0023], [0040], lines 1-5 and 12-16 of [0069], [0207], [0208], [0210], [0067], [0158] and lines 1-10 of [0009]); and

associating the applications with the assigned subsets of resources ([0023], [0040], [0177]).

27. In summary of the above citations, Hill teaches a networking system for allocating a plurality of resources/machines to a plurality of application programs in an optimized manner and reduces communication delay. The system determines resource availability of each of the plurality of machines in the network. The resource requirements of the applications are also collected such that it can be determined which resource or resources of the plurality/set of resources (assigned subset of available resources) gets allocated to the applications in conformance with considering the network capacity and the network bandwidth requirements of each application.

28. Hill does not expressly disclose that its application resource optimization ([0093]) is based on an objective function to reduce a number of network hops between processing resources in the assigned subset.

29. However, Lee teaches allocating network resources along an optimal path, wherein the path selection within routing is typically formulated as a shortest path optimization problem, i.e., by determining a series of network links connecting the source and destination such that a particular objective function is minimized. The objective function may be the number of hops, cost, delay, or some other metric (see Abstract; page 46, 1st two paragraphs). Hill and Lee are

Art Unit: 2195

analogous art because they are both in the same field of endeavor of network resource allocation and both are attempting to optimize its resource allocation such that delay is reduced. Thus, one of ordinary skill in the art would have known to modify Hill's network resource allocation system such that it would utilize an objective function to reduce a number of network hops between processing resources in the assigned subset, as taught in the reference of Lee. The suggestion/motivation for doing so would have been to provide an effective scheme to specifically calculate and determine the optimization of a path involving the minimal number of hops, delays, costs, etc. (see Abstract; page 46, 1st two paragraphs). Therefore, it would have been obvious to one of ordinary skill in the art to combine Hill and Lee to obtain the invention of claim 1.

30. As to claim 4, Hill teaches wherein the networking resources comprise network switches ([0031] and lines 4-7 of [0156]).

31. As to claim 9, it is rejected for the same reasons as stated in the rejections of claim 1.

32. As to claim 10, it is rejected for the same reasons as stated in the rejections of claim 1. In addition, Hill teaches a computer-readable medium configured with instructions for causing a processor of data processing arrangement to allocate resources to a plurality of applications ([0049]).

Art Unit: 2195

33. As to claims 11-13, they are rejected for the same reasons as stated in the rejections of claims 2-4, respectively.

34. As to claim 18, it is rejected for the same reasons as stated in the rejections of claim 1. In addition, Hill teaches a plurality of storage resources as part of the resources being allocated to the plurality of applications and network-coupled to the processing resources (Fig. 1, items 101, 102, 103, 115, 116, 111).

35. As to claim 19, it is rejected for the same reasons as stated in the rejections of claim 3.

36. Claims 5-6, 14-15, and 20-21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hill et al. (hereinafter Hill) (US 2004/0267897 A1) in view of Lee (“Routing Subject to Quality of Service Constraints in Integrated Communication Networks”, July/August 1995), and further in view of Varanasi et al. (hereinafter Varanasi) (US 7,443,799 B2).

37. As to claims 5-6, Hill discloses storage resources such as fixed storage 116, removable storage 115, ROM 103, etc., for a computer that is connected to a network with other computers

Art Unit: 2195

(Fig. 1). However, Hill in view of Lee is explicitly silent in teaching wherein the storage resources comprise a storage area network (SAN), wherein the storage area network includes at least one pair of redundant core switches coupled to storage devices, the core switches coupled to the processing resources via a plurality of edge switches. Varanasi discloses a networking system that can route data such as client/server applications to a plurality of resource devices, wherein the system is well suited to include a topology such as a SAN with at least one pair of redundant core switches 570 coupled to processing device resources 590 via a plurality of edge switches 580 (col. 1, lines 19-41 and 55-67, col. 2, lines 1-7 and 48-53, col. 8, lines 57-67 through col. 9, lines 1-8, Fig. 5). Hill, Lee and Varanasi are analogous art because they both are in the same field of endeavor of a network communication system that allocates between a plurality of applications and a plurality of resources. One of ordinary skill in the art would have known to modify Hill in view of Lee's network communication system such that it would include a SAN, wherein the storage area network includes at least one pair of redundant core switches coupled to storage devices, the core switches coupled to the processing resources via a plurality of edge switches, as taught in Varanasi's network communication system. By definition, a SAN is an architecture to attach remote computer storage devices to servers in such a way that the devices appear as locally attached to the operating system. Sharing storage usually simplifies storage administration and adds flexibility since cables and storage devices do not have to be physically moved to shift storage from one server to another. Other benefits include the ability to allow servers to boot from the SAN itself. This also allows for a quick and easy replacement of faulty servers, etc. Therefore, it would have been obvious to one of ordinary skill in the art to combine Hill, Lee, and Varanasi to obtain the invention of claims 5-6.

38. As to claims 14-15, they are rejected for the same reasons as stated in the rejections of claims 5-6, respectively.

39. As to claims 20-21, they are rejected for the same reasons as stated in the rejections of claims 5-6, respectively.

40. Claims 7-8, 16-17, and 22-23 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hill et al. (hereinafter Hill) (US 2004/0267897 A1) in view of Lee ("Routing Subject to Quality of Service Constraints in Integrated Communication Networks", July/August 1995), and further in view of Das et al. (hereinafter Das) (US 2005/0172291 A1).

41. As to claim 7, Hill in view of Lee is silent wherein reducing the communications delay between resources comprises solving a mixed-integer programming problem. However, Das teaches a dynamic resource allocation system that allocates resources amongst a plurality of application entities such that mixed-integer programming is used for optimization (Abstract, lines 5-21 of [0028]). One of ordinary skill in the art would have known to modify Hill in view of Lee's resource allocation system such that it would include mixed-integer programming, as taught in Das's resource allocation system. The suggestion/motivation for doing so would have

Art Unit: 2195

been to provide the predicted result of a dynamic, rapid, and optimal resource allocation in an automated fashion ([0002], [0005], and lines 5-21 of [0028]). Therefore, it would have been obvious to one of ordinary skill in the art to combine Hill, Lee, and Das to obtain the invention of claim 7.

42. As to claim 8, Hill ([0031], [0207], [0208], [0210], [0067]) and Das (Abstract, lines 5-21 of [0028]) teaches wherein the available resources include network switches coupled with the available resources, and the mixed-integer programming problem reduces communication delays between resources of the subset of the available resources by reducing data traffic on network links that interconnect the network switches.

43. As to claims 16-17, they are rejected for the same reasons as stated in the rejections of claims 7-8, respectively.

44. As to claims 22-23, they are rejected for the same reasons as stated in the rejections of claims 7-8, respectively.

45. Claims 24-27 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hill et al. (hereinafter Hill) (US 2004/0267897 A1) in view of Lee ("Routing Subject to Quality of

Art Unit: 2195

Service Constraints in Integrated Communication Networks”, July/August 1995), and further in view of Husain et al. (hereinafter Husain) (US 2003/0126260 A1).

46. As to claim 24, Hill in view of Lee do not expressly teach wherein the required resources of each application is specified in resource requirements that include attributes of the processing resources, wherein the attributes specify processor type and processor speed. However, Husain teaches a distributed resource managing system wherein the required resources of each application is specified in resource requirements that include attributes of the processing resources, wherein the attributes specify processor type and processor speed ([0011]; see claim 3; Abstract). One of ordinary skill in the art would have known to modify Hill in view of Lee with the teachings of Husain, as they all are in the same field of endeavor of network resource management. The suggestion/motivation for doing so would have been to provide the predicted result of improving speed, reducing error, and thus increasing the reliability of information for system ([0009]). Therefore, it would have been obvious to one of ordinary skill in the art to combine Hill, Lee, and Husain to obtain the invention of claim 24.

47. As to claim 25, Husain teaches wherein the resource requirements further specify storage patterns of files for each application, wherein determining the assigned subset is based on the resource requirements ([0072]).

48. As to claim 26, it is rejected for the same reasons as stated in the rejection of claim 24.

Art Unit: 2195

49. As to claim 27, it is rejected for the same reasons as stated in the rejection of claim 25.

50. Claims 1, 4, 9-10 and 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hill et al. (hereinafter Hill) (US 2004/0267897 A1) in view of Zhu et al. (US 2003/0120780 A1).

51. As to claim 1, Hill teaches a method for allocating resources to a plurality of applications (see Abstract), comprising:

determining, by a computer, available resources of a networked computing system, wherein the available resources comprise processing resources, networking resources, and storage resources (lines 7-8 of [0023]; Fig. 1, items 101, 102, 103, 115, 116, 111; lines 3-6 of paragraph [0097]);

determining, by the computer, for each application, required resources of the application (lines 9-17 of [0023]);

determining, by the computer, an assigned subset of the available resources for each application as a function of the required resources of the application and the available resources, wherein the function reduces communication delays between resources of the subset of the available resources in conformance with bandwidth capacity requirements of the application and in conformance with network bandwidth limitations ([0023], [0040], lines 1-5 and 12-16 of [0069], [0207], [0208], [0210], [0067], [0158] and lines 1-10 of [0009]); and

Art Unit: 2195

associating the applications with the assigned subsets of resources ([0023], [0040], [0177]).

52. In summary of the above citations, Hill teaches a networking system for allocating a plurality of resources/machines to a plurality of application programs in an optimized manner and reduces communication delay. The system determines resource availability of each of the plurality of machines in the network. The resource requirements of the applications are also collected such that it can be determined which resource or resources of the plurality/set of resources (assigned subset of available resources) gets allocated to the applications in conformance with considering the network capacity and the network bandwidth requirements of each application.

53. Hill does not expressly disclose that its application resource optimization ([0093]) is based on an objective function to reduce a number of network hops between processing resources in the assigned subset.

54. However, Zhu teaches network resource assignment system and method that optimizes the assignment of resources with respect to application requirements and desired objectives to reduce the number of network hops between processing resources in the assigned subset (see Abstract; [0036]; [0008]; [0021])

55. Hill and Zhu are analogous art because they are both in the same field of endeavor of network resource allocation and both are attempting to optimize its resource allocation such that delay is reduced. Thus, one of ordinary skill in the art would have known to modify Hill's network resource allocation system such that it would utilize an objective function to reduce a

Art Unit: 2195

number of network hops between processing resources in the assigned subset, as taught in the reference of Zhu. The suggestion/motivation for doing so would have been to provide an effective scheme to specifically calculate and determine the optimization of a path involving the minimal number of hops, delays, costs, etc. (see Abstract; page 46, 1st two paragraphs).

Therefore, it would have been obvious to one of ordinary skill in the art to combine Hill and Lee to obtain the invention of claim 1.

56. As to claims 4, 9-10 and 18, they are rejected for the same reasons as stated in the rejection of claim 1.

Response to Arguments

57. *Applicant argues that the newly amended claim limitations are not taught or suggested by the current prior art.*

In response, the applicant's amended claim prompted the new grounds of rejection, which render the argument moot.

Conclusion

The prior art made of record and not relied upon is considered pertinent to applicant's disclosure:

- **Kodialam et al. ("Minimum Interference Routing with Applications to MPLS Traffic Engineering", 2000)** teaches an objective function to reduce a number of

Art Unit: 2195

network hops between processing resources in the assigned subset (min-hop routing, etc.) (see Abstract).

- **Lee et al. ("Explicit routing with QoS constraints in IP over WDM, April 2002)** teaches an objective function to reduce a number of network hops between processing resources in the assigned subset (see Abstract).
- **Weaver (US 6,574,669 B1)** teaches a network resource assignment system and method that uses linear programming techniques with bandwidth constraints to reduce the number of hops (see Abstract; col. 1, lines 13-64; col. 2, lines 1-25).

Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a).

Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire **THREE MONTHS** from the mailing date of this action. In the event a first reply is filed within **TWO MONTHS** of the mailing date of this final action and the advisory action is not mailed until after the end of the **THREE-MONTH** shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than **SIX MONTHS** from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to **KENNETH TANG** whose telephone number is (571)272-3772.

The examiner can normally be reached on 8:30AM - 6:00PM, Every other Friday off.

Art Unit: 2195

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Meng-Ai An can be reached on (571) 272-3756. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Meng-Ai An/
Supervisory Patent Examiner, Art Unit 2195

/Kenneth Tang/
Examiner, Art Unit 2195